

# An Improving EFA for Clutter Suppression by Using the Persymmetric Covariance Matrix Estimation

Yan Zhou<sup>1</sup> · Lin Wang<sup>1</sup> · Xiaoxuan Chen<sup>1</sup> ·  
Cai Wen<sup>1</sup> · Bo Jiang<sup>1</sup> · Dingyi Fang<sup>1</sup>

Received: 13 July 2017 / Revised: 25 December 2017 / Accepted: 26 December 2017 /  
Published online: 3 January 2018  
© Springer Science+Business Media, LLC, part of Springer Nature 2018

**Abstract** The extended factored approach (EFA) is believed to be one of the most efficient and practical space–time adaptive processing (STAP) algorithms for clutter suppression in an airborne radar system. However, it cannot effectively work in the airborne radar system with large antenna array for the huge computational cost and the lack of training sample. To solve these problems, a bi-iterative algorithm based on the persymmetric covariance matrix estimation is proposed in this paper. Firstly, the clutter covariance matrix is estimated by using the original data, the constructed spatial transformed data, the constructed temporal transformed data and the constructed spatial–temporal transformed data. Secondly, the spatial weight vector in EFA is decomposed as the Kronecker products of two short weight vectors. Finally, the bi-iterative algorithm is exploited to obtain the desired weight vectors. Thus, the improving EFA with small training sample demanding is realized. Experimental results

---

✉ Yan Zhou  
yanzhou@nwu.edu.cn

Lin Wang  
wanglin@nwu.edu.cn

Xiaoxuan Chen  
dadayuasi@163.com

Cai Wen  
20165053@nwu.edu.cn

Bo Jiang  
bojiangjiang@126.com

Dingyi Fang  
dyf@nwu.edu.cn

<sup>1</sup> School of Information Science and Technology, Northwest University, Xi'an 710127, China

demonstrate the effectiveness of the proposed method under small training sample support.

**Keywords** Space–time adaptive processing (STAP) · Clutter suppression · Bi-iterative · Airborne radar

## 1 Introduction

Space–time adaptive processing (STAP) has been viewed as an effective tool in an airborne surveillance radar since it was first proposed by Brennan et al. [1]. It adaptively suppresses the clutter received by the airborne radar from both the temporal domain and spatial domain and also detects moving targets that are relatively weak compared with the surrounding clutter. One of the fundamental problems in STAP is the estimation of the clutter covariance matrix (CCM). As stated in [8], the sample covariance matrix (SCM) is a practical estimator to compute CCM. However, in the full-dimension STAP (f-STAP), SCM requires excessive number of training samples, which is difficult to be realized in the real airborne radar environment. Moreover, f-STAP needs to conduct matrix inverse operation in adaptive processing, which is quite computationally intensive and intractable. These two main limitations motivate the development of STAP algorithms with low computational cost and small training sample demanding. Sub-optimal dimension/rank-reduced STAP algorithms, which maintain nearly the same performance as f-STAP and address the above issues, have been studied in various literature [2–4, 6, 7, 9, 12–15, 17, 18]. Among them, the post-Doppler adaptive processing method, named extended factored approach (EFA or mDT), which is expected to be one of the most efficient and practical STAP algorithms, is mainly concerned in this paper [3, 16].

The basic principle of EFA is that it transforms f-STAP of a  $NK$ -dimensional adaptive filtering problem into  $K$  separate  $PN$ -dimensional adaptive processing problems, where  $N$  is the number of antennas,  $K$  is the number of pulses, and  $P$  is an integer satisfied  $P > 1$ . As a consequence, the required number of homogeneous training samples is reduced to  $2PN$  according to the RMB rule [10]. Besides, the computational cost of EFA is also decreased and the applicability of it is increased in practice. Unfortunately, insufficient homogeneous training sample support is a frequent condition in the airborne radar system with large-scale antenna array elements, which results in considerable degeneration of the clutter suppression ability of EFA.

Actually, nowadays the persymmetry property of CCM is usually used as the a priori knowledge to increase the training sample support for accurate CCM estimation. It is initially exploited for communications. Tong et al. [12] extended the application of persymmetry to EFA (Per-EFA) and showed that the required number of training samples is reduced. Per-EFA is demonstrated to be an effective algorithm in training-limited scenarios. However, the computational cost of Per-EFA in adaptive processing increases because of the extra computational cost in CCM estimation. Besides, it still requires excessive quantity of training samples in the airborne radar system with large spatial degrees of freedom (DoFs).

In order to reduce both the computational cost and the training sample requirement in the post-Doppler adaptive processing, a bi-iterative algorithm based on Per-EFA is proposed in this paper. Firstly, according to the persymmetric property, CCM is estimated by using the original data, the constructed spatial transformed data, the constructed temporal transformed data and the constructed spatial–temporal transformed data. Secondly, the spatial weight vector in EFA is decomposed as the Kronecker products of two short weight vectors. Then the cost function of EFA is transformed into the cost function with multiple weight vectors. Finally, the bi-iterative algorithm [11] is exploited to obtain the desired weight vectors.

**Notation** Throughout this paper, matrix and vector are represented by boldface uppercase letter and boldface lowercase letter, respectively. Superscripts  $T$ ,  $*$  and  $H$  separately denote transpose, conjugate and conjugate transpose of a matrix or a vector.  $\otimes$  and  $\circ$  denote the Kronecker product operator and Hadamard product operator, respectively. A  $c \times c$  identity matrix is represented as  $\mathbf{I}_c$ .  $\|\bullet\|_F$  denotes the F-norm of a matrix, and  $[\bullet]^{-1}$  means the inverse of a matrix.

## 2 STAP Fundamentals

In this paper, we assume that the uniform linear antenna array is paralleled to the ground. The configuration of a traditional airborne phased-array radar system to be studied is shown in Fig. 1. The airborne radar system moves at a constant speed  $v_a$  along the positive x-direction with  $N$  antenna array elements and processes  $K$  pulses on the repetition of  $T_r$  in one coherent processing interval (CPI). Let the crab angle between the antenna array line and the flight velocity be  $\theta_c$ . The azimuth angle and elevation angle at the  $i$ th clutter scatterer in the  $l$ th range cell are denoted as  $\theta_i$  and  $\varphi_l$ , respectively.

For a given range cell, the clutter, generated by the terrain scatterings, can be considered as a collection of independent scatters as [8]

$$\mathbf{z}(l) = \sum_{i=1}^{N_c} \beta_i \mathbf{a}_t(\theta_i, \varphi_l) \otimes \mathbf{a}_s(\theta_i, \varphi_l), \quad (1)$$

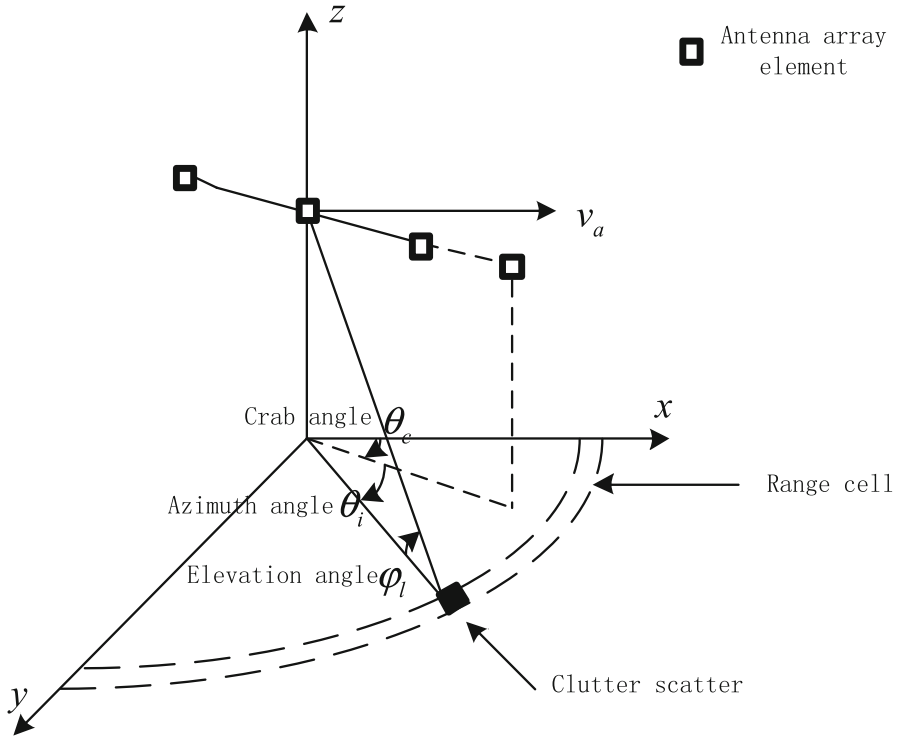
where  $N_c$  scatters are divided in an iso-range cell,  $\beta_i$  is clutter echo's random complex amplitude,  $\mathbf{a}_t(\theta_i, \varphi_l)$  and  $\mathbf{a}_s(\theta_i, \varphi_l)$  are temporal steering-vector and spatial steering-vector, respectively. They can be expressed as

$$\mathbf{a}_t(\theta_i, \varphi_l) = [1, \exp(j2\pi\omega_i), \dots, \exp(j2\pi(K-1)\omega_i)]^T, \quad (2)$$

$$\mathbf{a}_s(\theta_i, \varphi_l) = [1, \exp(j2\pi\nu_i), \dots, \exp(j2\pi(N-1)\nu_i)]^T, \quad (3)$$

where  $\omega_i = 2T_r v_a \cos(\theta_i + \theta_c) \cos \varphi_l / \lambda$  is the normalized Doppler frequency,  $\nu_i = d \cos \theta_i \cos \varphi_l / \lambda$  is the normalized spatial frequency,  $\lambda$  is the operating wavelength, and  $d$  is the array element spacing. The received clutter signal also contains noise, namely

$$\mathbf{x}(l) = \mathbf{z}(l) + \mathbf{n}, \quad (4)$$



**Fig. 1** The airborne phased-array radar system

where  $\mathbf{n}$  is the Gaussian noise vector. Let  $\mathbf{s} \in \mathbf{C}^{KN \times 1}$  denote the normalized target vector. The objective of f-STAP is to maintain the output energy of the target while minimizing that of clutter. This implies the following cost function

$$\begin{cases} \min E \left[ |\mathbf{w}^H \mathbf{x}(l)|^2 \right] \\ \text{s.t. } \mathbf{w}^H \mathbf{s} = 1 \end{cases} \tag{5}$$

The Lagrange multiplier methodology is adopted to obtain the optimal weight

$$\mathbf{w} = \mathbf{R}_x^{-1} \mathbf{s} / \left( \mathbf{s}^H \mathbf{R}_x^{-1} \mathbf{s} \right), \tag{6}$$

where  $\mathbf{R}_x = E[\mathbf{x}(l)\mathbf{x}^H(l)]$  is the CCM. In general, since the statistics of the range cell under test is never known, CCM is replaced by the SCM estimator, namely  $\tilde{\mathbf{R}}_x = \frac{1}{L} \sum_{l=1}^L \mathbf{x}(l)\mathbf{x}^H(l)$ , where  $L$  is the number of training samples obtained in adjacent range cells. RMB rule shows that, in the Gaussian noise environment, if the number of homogeneous training samples exceeds twice the dimension of CCM, the output SCNR (signal-to-clutter-plus-noise ratio) loss will be within  $3dB$ . The training sample requirement of f-STAP is evidently unpractical. Moreover, as the inverse operation of a  $KN \times KN$  matrix, the computational cost of f-STAP is unbearable of approximate

$O(K^3N^3)$ . These two major issues prevent f-STAP from being put into effect. Hence, EFA, one of the most efficient and practical STAP algorithms, is mainly concerned in this paper. It performs Doppler filtering on data of each array element and then spatial filtering separately within  $P$  ( $P \geq 2$ ) Doppler bins. Generally,  $P$  is set to 3 in EFA. Therefore, the EFA with three Doppler bins is referred in the following. Let  $\mathbf{f}_k = \exp(-j2\pi(k-1)[0, 1, 2, \dots, K-1]^T/K)/\text{sqrt}(K)$ ,  $k = 1, \dots, K$ , which are actually the coefficients of discrete Fourier transform and compose the Doppler filter group, be the Doppler filter coefficient vector. Then the EFA dimension-reduced transformation matrix according to the  $k$ th Doppler bin can be expressed as

$$\mathbf{B}_k = [\mathbf{f}_{k-1}, \mathbf{f}_k, \mathbf{f}_{k+1}]^T \otimes \mathbf{I}_N. \quad (7)$$

The dimension-reduced space–time data and target can be achieved as

$$\begin{aligned} \hat{\mathbf{x}}_k(l) &= \mathbf{B}_k \mathbf{x}(l) \\ \hat{\mathbf{s}}_k &= \mathbf{B}_k \mathbf{s} \end{aligned}. \quad (8)$$

The cost function in EFA is

$$\begin{cases} \min E \left[ |\hat{\mathbf{w}}_k^H \hat{\mathbf{x}}_k(l)|^2 \right] \\ \text{s.t. } \hat{\mathbf{w}}_k^H \hat{\mathbf{s}}_k = 1 \end{cases}. \quad (9)$$

Similarly, the solution to Eq. (9) is obtained by using the Lagrange multiplier methodology and the CCM is estimated by

$$\hat{\mathbf{R}}_{\hat{\mathbf{x}}_k} = \frac{1}{L} \sum_{l=1}^L \hat{\mathbf{x}}_k(l) \hat{\mathbf{x}}_k^H(l). \quad (10)$$

Evidently, the dimension of CCM is reduced to  $3N$ . As a consequence, the required number of training samples and computational cost in adaptive processing are separately reduced to  $6N$  and  $O(3^3N^3)$ . Although EFA reduces the training sample requirement, its clutter suppression ability can still be heavily degraded in large-scale antenna array elements system.

### 3 The Bi-iterative Algorithm Based on Per-EFA

Through the persymmetric property of CCM, Ref. [12] increases the number of training samples by constructing the spatial transformed data, the temporal transformed data and spatial–temporal transformed data. Let  $f_{k-1}$ ,  $f_k$ ,  $f_{k+1}$  be the normalized Doppler frequency of three adjacent Doppler bins,  $\mathbf{U}_t = \text{diag}([e^{j2\pi f_{k-1}(N-1)}, e^{j2\pi f_k(N-1)}, e^{j2\pi f_{k+1}(N-1)}])$ ,  $\mathbf{U}_{st} = \text{diag}([e^{-j2\pi f_{k-1}(N-1)}, e^{-j2\pi f_k(N-1)}, e^{-j2\pi f_{k+1}(N-1)}])$  and  $\mathbf{J}_s = \text{fliplr}(\mathbf{I}_N)$  with  $\text{fliplr}(\bullet)$  being the MATLAB operator that flips matrix in left/right direction. Then CCM can be estimated by using the constructed data as

$$\widetilde{\mathbf{R}}_{\hat{x}_k} = \left( \hat{\mathbf{R}}_{\hat{x}_k} + \mathbf{B}_s \hat{\mathbf{R}}_{\hat{x}_k}^* \mathbf{B}_s^H + \mathbf{B}_t \hat{\mathbf{R}}_{\hat{x}_k} \mathbf{B}_t^H + \mathbf{B}_{st} \hat{\mathbf{R}}_{\hat{x}_k}^* \mathbf{B}_{st}^H \right) / 4, \tag{11}$$

where

$$\mathbf{B}_s = \mathbf{I}_p \otimes \mathbf{J}_s, \quad \mathbf{B}_t = \mathbf{U}_t \otimes \mathbf{I}_N, \quad \mathbf{B}_{st} = \mathbf{U}_{st} \otimes \mathbf{J}_s, \tag{12}$$

in which  $\mathbf{B}_s$  is the spatial transformation matrix that exploits the persymmetric property of CCM in the spatial domain,  $\mathbf{B}_t$  is the temporal transformation matrix that exploits the persymmetric property of CCM in the temporal domain, and  $\mathbf{B}_{st}$  is the spatial–temporal transformation matrix that jointly exploits the persymmetric property of CCM in the spatial and temporal domains.

Then, the proposed bi-iterative adaptive method will be introduced after constructing the CCM estimator by Eq. (11). As a matter of fact,  $\hat{\mathbf{x}}_k(l) = [\hat{x}_1(l), \hat{x}_2(l), \dots, \hat{x}_{PN}(l)]$  can be rearranged as the matrix with the following formula

$$\hat{\mathbf{X}}_k(l) = \begin{bmatrix} \hat{X}_{1,1}(l), & \hat{X}_{1,2}(l), & \dots, & \hat{X}_{1,b}(l) \\ \hat{X}_{2,1}(l), & \hat{X}_{2,2}(l), & \dots, & \hat{X}_{2,b}(l) \\ \vdots & \vdots & \vdots & \vdots \\ \hat{X}_{a,1}(l), & \hat{X}_{a,2}(l), & \dots, & \hat{X}_{a,b}(l) \end{bmatrix}, \tag{13}$$

where  $\hat{X}_{n,m}(l) = \hat{x}_{(n-1) \times b + m}(l), n = 1, \dots, a, m = 1, \dots, b$  and  $a \times b = P \times N$  with scalars  $a$  and  $b$  both being integers. The objective function in Eq. (9) can be transformed into the following formula by using Eq. (13)

$$E \left[ \left| \hat{\mathbf{w}}_k^H \hat{\mathbf{x}}_k(l) \right|^2 \right] = E \left[ \left\| \hat{\mathbf{W}}_k^* \circ \hat{\mathbf{X}}_k(l) \right\|_F^2 \right] = E \left[ \left| \sum_{m=1}^b \sum_{n=1}^a \hat{w}_{n,m}^* \hat{X}_{n,m}(l) \right|^2 \right]. \tag{14}$$

According to the separation of variables, let weight coefficient  $\hat{w}_{n,m} = u_n v_m$  be expressed as the separable form and be introduced into Eq. (14), and the following bi-quadratic cost function can be acquired

$$\begin{aligned} E \left[ \left| \sum_{m=1}^b \sum_{n=1}^a \hat{w}_{n,m}^* \hat{X}_{n,m}(l) \right|^2 \right] &= E \left[ \left| \sum_{m=1}^b \sum_{n=1}^a u_n^* \hat{X}_{n,m}(l) v_m^* \right|^2 \right] \\ &= E \left[ \left| \mathbf{u}^H \hat{\mathbf{X}}_k(l) \mathbf{v}^* \right|^2 \right] = E \left[ \left| (\mathbf{u} \otimes \mathbf{v})^H \hat{\mathbf{x}}_k(l) \right|^2 \right], \end{aligned} \tag{15}$$

where  $\mathbf{u} \in \mathbf{C}^{a \times 1}$  and  $\mathbf{v} \in \mathbf{C}^{b \times 1}$  represent the separable weight vectors. Therefore, after substituting  $\hat{\mathbf{R}}_{\hat{x}_k}$  by  $\widetilde{\mathbf{R}}_{\hat{x}_k}$ , Eq. (9) can be converted into the following cost function

$$\begin{cases} \min (\mathbf{u} \otimes \mathbf{v})^H \widetilde{\mathbf{R}}_{\hat{x}_k} (\mathbf{u} \otimes \mathbf{v}) \\ \text{s.t. } (\mathbf{u} \otimes \mathbf{v})^H \hat{\mathbf{s}}_k = 1 \end{cases}. \tag{16}$$

Equation (16) means to minimize the output energy of the clutter while maintain that of the target by jointly optimizing the weight vectors  $\mathbf{u}$  and  $\mathbf{v}$ . By using the Lagrange multiplier methodology, Eq. (16) can be transformed into the following unconstrained one

$$J(\mathbf{u}, \mathbf{v}) = (\mathbf{u} \otimes \mathbf{v})^H \widetilde{\mathbf{R}}_{\hat{\mathbf{x}}_k} (\mathbf{u} \otimes \mathbf{v}) + \mu(1 - (\mathbf{u} \otimes \mathbf{v})^H \hat{\mathbf{s}}_k), \quad (17)$$

where  $\mu$  is the Lagrange multiplier. Generally, the solution to Eq. (17) is not easy to be analytically expressed. Fortunately, the bi-iterative algorithm [11] can be conveniently applied to numerically solve this cost function. And the procedure of the algorithm is as follows.

To obtain the numerical solution to Eq. (17),  $\mathbf{v}$  is initialized with  $\mathbf{v}(0)$ . Then Eq. (17) becomes an unconstrained cost function with respect to  $\mathbf{u}$  and  $\mu$  after substituting  $\mathbf{v}(0)$  into it. Let the gradient of  $J(\mathbf{u}, \mathbf{v}(0))$  with respect to  $\mathbf{u}$  and  $\mu$  be zero, the following iterative value can be obtained

$$\mathbf{u}(1) = \mathbf{R}_u^{-1} \mathbf{s}_u / \left( \mathbf{s}_u^H \mathbf{R}_u^{-1} \mathbf{s}_u \right), \quad (18)$$

where

$$\mathbf{R}_u = (\mathbf{I}_a \otimes \mathbf{v}(0))^H \widetilde{\mathbf{R}}_{\hat{\mathbf{x}}_k} (\mathbf{I}_a \otimes \mathbf{v}(0)), \quad \mathbf{s}_u = (\mathbf{I}_a \otimes \mathbf{v}(0))^H \hat{\mathbf{s}}_k. \quad (19)$$

Again, with fixing  $\mathbf{u}(1)$ , let the gradient of  $J(\mathbf{u}(1), \mathbf{v})$  with respect to  $\mathbf{v}$  and  $\mu$  be zero, and the iterative value of  $\mathbf{v}(1)$  can be achieved by

$$\mathbf{v}(1) = \mathbf{R}_v^{-1} \mathbf{s}_v / \left( \mathbf{s}_v^H \mathbf{R}_v^{-1} \mathbf{s}_v \right), \quad (20)$$

where

$$\mathbf{R}_v = (\mathbf{u}(1) \otimes \mathbf{I}_b)^H \widetilde{\mathbf{R}}_{\hat{\mathbf{x}}_k} (\mathbf{u}(1) \otimes \mathbf{I}_b), \quad \mathbf{s}_v = (\mathbf{u}(1) \otimes \mathbf{I}_b)^H \hat{\mathbf{s}}_k. \quad (21)$$

By repeating the above iterative procedure,  $\mathbf{u}(2)$  and  $\mathbf{v}(2)$ ,  $\mathbf{u}(3)$  and  $\mathbf{v}(3)$ , ..., can be obtained in turn. The iterative procedure will be stopped if the termination condition  $\|\mathbf{v}(t) - \mathbf{v}(t-1)\| / \|\mathbf{v}(t)\| < \delta$  is satisfied at the  $t$ th iteration, where  $\delta$  is the threshold and  $\|\bullet\|$  is the Euclidean norm of a vector. Finally,  $\hat{\mathbf{w}}_k$  is achieved by the Kronecker products of two short vectors

$$\hat{\mathbf{w}}_k = \mathbf{u}(t) \otimes \mathbf{v}(t). \quad (22)$$

The dimensions of  $\mathbf{R}_u$  and  $\mathbf{R}_v$  produced in the iterative procedure are  $a \times a$  and  $b \times b$ , respectively. After weight decomposition, both the required number of training samples and computational cost in the spatial adaptive processing are reduced. Furthermore, in order to satisfy the small computational cost and training sample

requirements, the constraints  $\min |a - b|$  and  $a < b$  are added. Hereafter, for simplicity, the bi-iterative algorithm based on the persymmetry of CCM is denoted as BiPer-EFA.

It is worth noting that  $J(\mathbf{u}, \mathbf{v}) = J(\alpha\mathbf{u}, \alpha^{-1}\mathbf{v})$  is satisfied for any nonzero constant  $\alpha$ . In order to eliminate the scale indeterminacy,  $\|\mathbf{v}\|$  is set to unity in each iteration procedure.

### 4 Convergence and Computational Cost Analysis

The convergence of the proposed bi-iterative algorithm is verified in this section. Firstly, let  $\|\mathbf{v}\| = 1$ , the relation  $J(\mathbf{u}(t), \mathbf{v}(t)) = \mathbf{u}^H \mathbf{R}_u \mathbf{u} \geq \kappa_{\min} \|\mathbf{u}\|^2$  can be acquired and the set  $\{\mathbf{u}(t) | \kappa_{\min} \|\mathbf{u}(t)\|^2 \leq J(\mathbf{u}(t), \mathbf{v}(t)) \leq c\}$  is bounded for any constant  $0 < c < \infty$ , where  $\kappa_{\min}$  is the smallest eigenvalue of  $\mathbf{R}_u$ . Therefore the set  $\{\mathbf{u}(t), \mathbf{v}(t) | J(\mathbf{u}(t), \mathbf{v}(t)) \leq c\}$ , which is bounded for any constant  $0 < c < \infty$  and iteration number  $t$ , can be inferred. Secondly, the cost function  $J(\mathbf{u}, \mathbf{v})$  in Eq. (17) is continuous for its differentiability. Thirdly,  $J(\mathbf{u}(t - 1), \mathbf{v}(t - 1)) \geq J(\mathbf{u}(t), \mathbf{v}(t))$  since

$$\begin{aligned} J(\mathbf{u}(t - 1), \mathbf{v}(t - 1)) &\geq J(\mathbf{u}(t - 1), \mathbf{v}(t)) = \min_{\mathbf{v}} J(\mathbf{u}(t - 1), \mathbf{v}) \\ &\geq J(\mathbf{u}(t), \mathbf{v}(t)) = \min_{\mathbf{u}} J(\mathbf{u}, \mathbf{v}(t)) \end{aligned} \tag{23}$$

The three points just mentioned above satisfy the definition of Lyapunov function [5, 7]. Accordingly,  $J(\mathbf{u}, \mathbf{v})$  is a Lyapunov function and the convergence of the bi-iterative algorithm is proved.

Since the computational costs of the Doppler filtering are nearly the same in each method, we mainly consider the computational cost of CCM estimation and matrix inversion in the spatial adaptive processing. The multiplication and division number (MDN) is used as an index of computational cost. First of all, the computational cost of BiPer-EFA is analyzed in detail. It should be noted that the MDN of reducing the data dimension and constructing the transformed data is  $L_1 [3N^2K + 3(3N)^2]$ . In the iteration algorithm, the MDN of CCM estimation and weight vector computation is  $12(a + b)NL_1 + \frac{2}{3}(a^3 + b^3) + 2(a^2 + b^2) + a + b$ . As shown in Fig. 5, the bi-iterative algorithm can reach the convergence value within 10 steps. Therefore, the total MDN of BiPer-EFA is

$$L_1 [3N^2K + 3(3N)^2] + 10 \left[ 12(a + b)NL_1 + \frac{2}{3}(a^3 + b^3) + 2(a^2 + b^2) + a + b \right]. \tag{24}$$

Similarly, the MDNs of EFA and Per-EFA are

$$L_3 [3N^2K + (3N)^2] + \frac{2}{3}(3N)^3 + 2(3N)^2 + 3N, \tag{25}$$

$$L_2 [3N^2K + 7(3N)^2] + \frac{2}{3}(3N)^3 + 2(3N)^2 + 3N, \tag{26}$$



respectively, where  $L_1$ ,  $L_2$  and  $L_3$  are the required number of training samples in each method.

## 5 Simulation Results

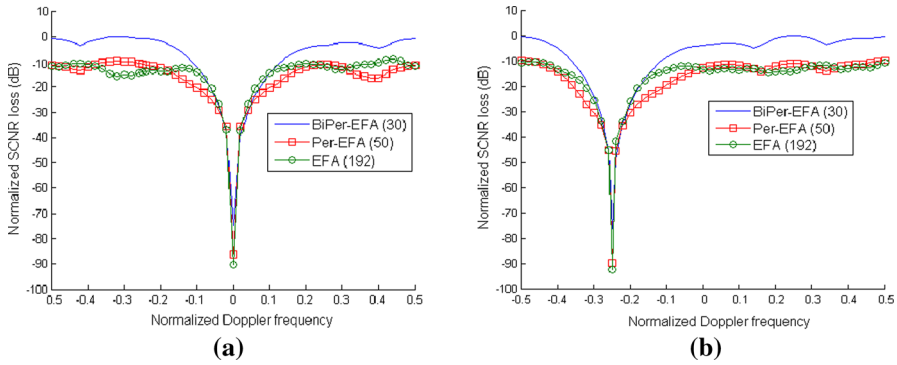
The linear array element mounted on the aircraft is composed of 64 antennas with equally spaced by a distance of  $d = 0.1$  m. The echo amplitude of each clutter source is a complex Gaussian random variable and has been weighted by the transmit beam pattern. The other parameters are listed in Table 1. The experimental platform is on a PC notebook with 2.4 GHz Core i7-5500U processor, 4 GB RAM. The algorithms are coded as MATLAB programs and are run in MATLAB R2011a.

One important metric to measure the clutter suppression ability of a STAP method is the SCNR loss, which is defined as the loss between the implemented processor and the optimal one [8]. According to the parameters given above, the weight vector as discussed in Eq. (15) is decomposed as the Kronecker products of two short vectors  $\mathbf{u} \in \mathbb{C}^{12 \times 1}$  and  $\mathbf{v} \in \mathbb{C}^{16 \times 1}$ . Figure 2 shows the normalized SCNR loss (normalized by the maximum SCNR loss among EFA, Per-EFA and BiPer-EFA) curves of EFA, Per-EFA and BiPer-EFA versus the normalized Doppler frequency with small training sample support. For the sake of estimated CCMs being invertible, the number of training samples separately conducted in each method is 30, 50 and 192, which are marked in the parentheses in Fig. 2. In addition,  $\theta_c = 0$  that represents the sidelooking scenario and  $\theta_c = \pi/6$  that represents the nonsidelooking scenario are considered in plotting the curves of Fig. 2. It is demonstrated that though less training samples are exploited in BiPer-EFA, it obviously possesses better SCNR loss performance than those of Per-EFA and EFA in both main clutter region and sidelobe clutter region. Figure 3 shows the normalized SCNR loss curves of EFA, Per-EFA and BiPer-EFA versus the normalized Doppler frequency with large training sample support ( $L = 500$ ) under sidelooking scenario and nonsidelooking scenario. Under this condition, the SCNR performance of BiPer-EFA is a slightly lower than those of Per-EFA and EFA. Evidently, when adequate number of training samples can be acquired, EFA possesses the best SCNR performance.

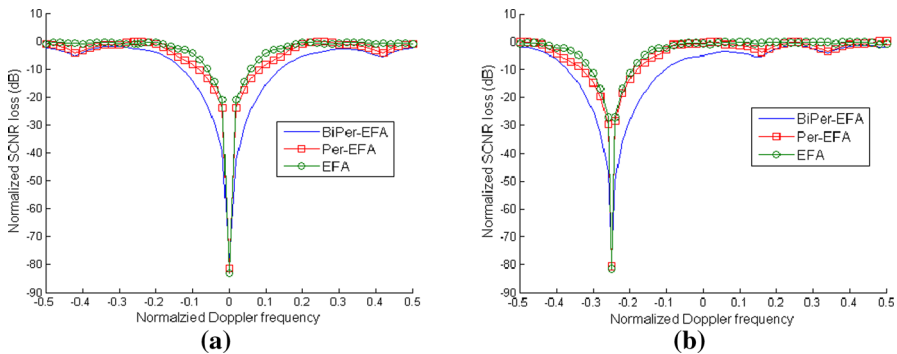
The SCNR loss curve, varying with the number of training samples, describes the training sample convergence rate of a STAP method. It also demonstrates the min-

**Table 1** Parameters of the simulation

Parameters	Value
Wavelength	0.2 m
Pulse repetition frequency (PRF)	2000 Hz
Number of pulses in a CPI	16
Platform velocity	100 m/s
Platform height	8 km
Azimuth angle of the main beam	$\pi/2$
Elevation angle of the main beam	0

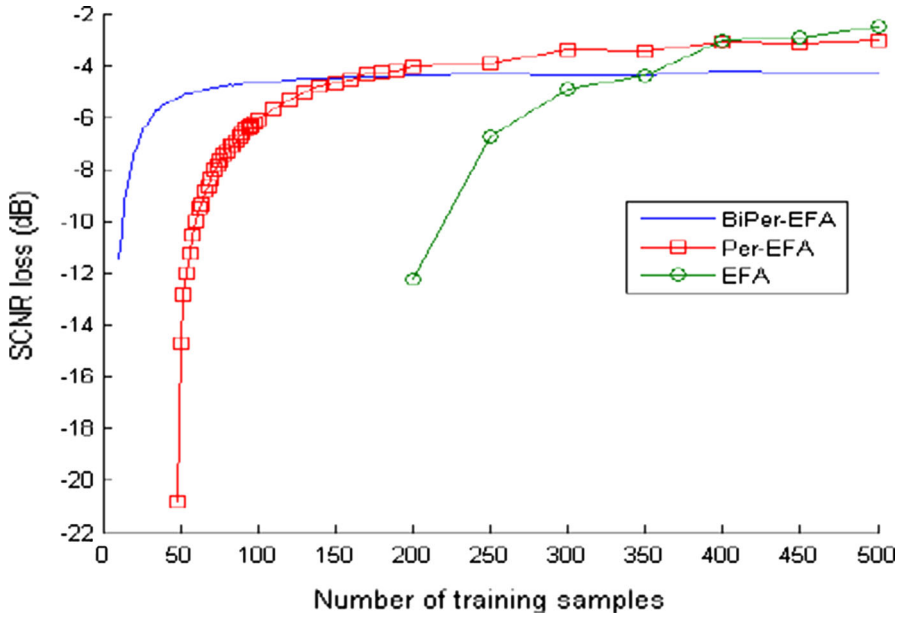


**Fig. 2** SCNR loss curves with small training sample support. **a**  $\theta_c = 0$ , **b**  $\theta_c = \pi/6$

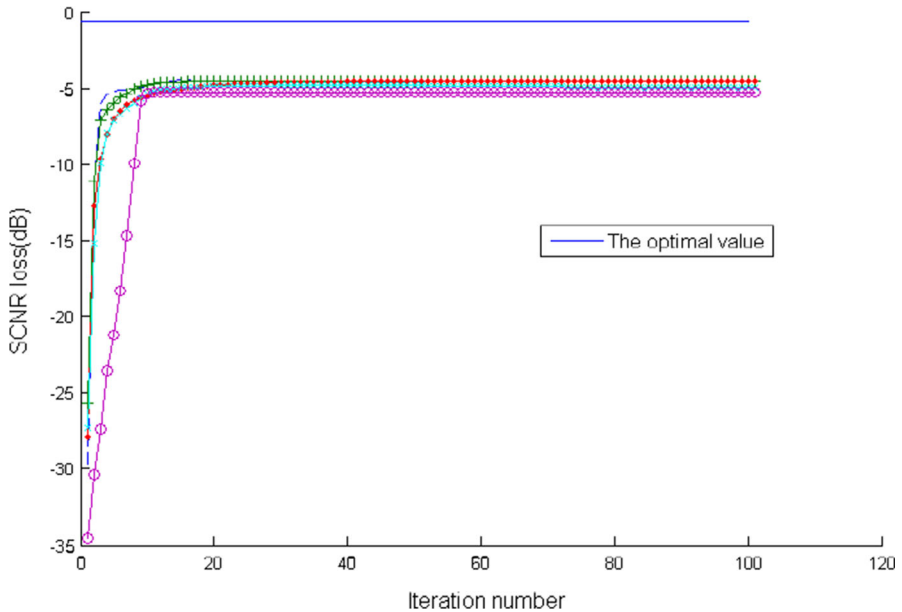


**Fig. 3** SCNR loss curves with large training sample support ( $L = 500$ ). **a**  $\theta_c = 0$ , **b**  $\theta_c = \pi/6$

imum number of training samples required in a STAP method. Figure 4 shows the convergence rates of BiPer-EFA, Per-EFA and EFA with the initial number of training samples individually being 10, 50 and 192 at normalized Doppler frequency 0.3 and normalized spatial frequency 0. Evidently, BiPer-EFA enjoys far faster training sample convergence rate than its counterparts. In other words, BiPer-EFA owns lower training sample requirement than Per-EFA and EFA. The experimental result of iterative convergences of BiPer-EFA with five different random initializations is shown in Fig. 5. The optimal value in Fig. 5 means the SCNR loss value obtained by the optimal spatial weight. The BiPer-EFA with all the five initializations can be convergent within 10 iterations, which demonstrates its fast iterative convergence rate. However, all the five convergent values are around  $-5$  dB, which probably means that the bi-iterative algorithm may be attracted by the local minimum. Besides, when the number of training sample is large enough, the SCNR loss performances of EFA and Per-EFA are superior to that of BiPer-EFA. The probable reason for this result is that when adequate number of training samples is available, Per-EFA and EFA can achieve the optimal SCNR values, while BiPer-EFA can only achieve an inferior SCNR value for the limitation of its special structure. Through the experimental results shown above, the advantage of BiPer-EFA over Per-EFA and EFA under small training sample support is exhibited.

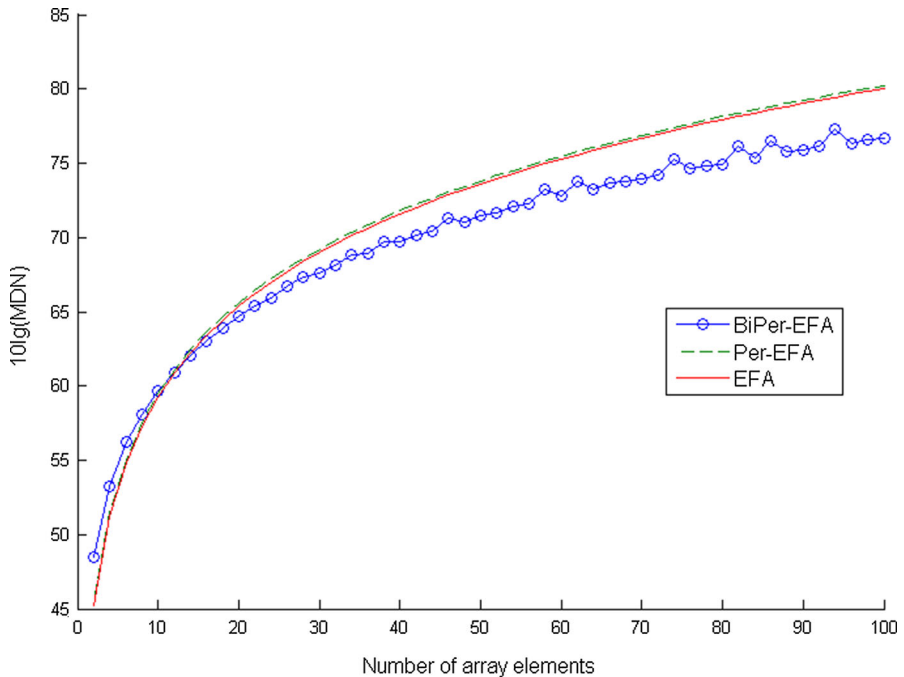


**Fig. 4** The comparison of convergence rates versus the number of training samples



**Fig. 5** The iterative convergence rate with five different initializations

According to Eqs. (24), (25) and (26), the computational cost comparison is shown in Fig. 6. Here,  $L_1 = 30$ ,  $L_2 = 100$  and  $L_3 = 250$  are adopted since each method



**Fig. 6** Computational cost comparison

**Table 2** Processing time comparison under small training sample support

BiPer-EFA	19.78 s
Per-EFA	23.94 s
EFA	23.21 s

can acquire nearly the same SCNR loss with the chosen number of training samples. Compared with EFA and Per-EFA, one conclusion is obtained that BiPer-EFA can achieve the smallest computational cost when  $N > 10$ . According to the simulation result, the computational cost of BiPer-EFA is almost 2 times smaller than that of Per-EFA and that of EFA when the array element scale approaches large. In addition, for the additional CCM estimation by Eq. (11), the computational cost of Per-EFA is a little higher than that of EFA.

The processing times for drawing the curves of each method in Fig. 2 are compared in Table 2. Obviously, in accordance with the aforementioned computational cost analysis, BiPer-EFA possesses the minimum processing time among all methods. In Table 3, the processing times for drawing the curves of each method in Fig. 3 are compared. When  $L_1 = L_2 = L_3 = 500$ , BiPer-EFA achieves the maximum computational cost, while EFA achieves the minimum computational cost. As a consequence, when adequate number of training samples can be obtained, EFA enjoys better clutter suppression ability and smaller computational cost.

**Table 3** Processing time comparison under large training sample support

BiPer-EFA	36.39 s
Per-EFA	35.28 s
EFA	30.54 s

## 6 Conclusion

In this paper, an algorithm that decreases the training sample requirement and computational cost has been proposed. The persymmetry property of CCM is utilized, and the weight vector of EFA is decomposed. Then, the bi-iterative algorithm is applied to find the desired weight vectors. Experimental results show that the proposed algorithm has superiority in training sample convergence rate and computational cost in the airborne radar system with large antenna array.

**Acknowledgements** The authors would like to thank very much the Handling Editor and the anonymous reviewers for their valuable comments and suggestions that have significantly improved the manuscript. This work was sponsored in part by National Natural Science Foundation of China under Grants 61503300 and 41601353 and the Scientific Research Plan of Education Department of Shaanxi Province under Grant 17JK0789.

## References

1. L.E. Brennan, I.S. Reed, Theory of adaptive radar. *IEEE Trans. Aerosp. Electron. Syst.* **9**(2), 237–252 (1973)
2. S. Chen, L.J. Kong, J.Y. Yang, Adaptive clutter nulling approach for heterogeneous environments. *Circuits Syst. Signal Process.* **34**(3), 987–1000 (2015)
3. R.C. Dipietro, Extended factored space–time processing for airborne radar systems, in *Signals, Systems and Computers, Conference Record of The Twenty-Sixth Asilomar Conference*, IEEE (1992), pp. 425–430
4. R. Fa, R.C. Lamare, L. Wang, Reduced-rank STAP schemes for airborne radar based on switched joint interpolation, decimation and filtering algorithm. *IEEE Trans. Signal Process.* **58**(8), 4182–4194 (2010)
5. D.Z. Feng, W.X. Zheng, A. Cichocki, Matrix-group algorithm via improved whitening process for extracting statistically independent sources from array signals. *IEEE Trans. Signal Process.* **55**(3), 962–977 (2007)
6. J.P. Lasalle, *The Stability of Dynamical Systems* (SIAM Press, Philadelphia, 1976)
7. N. Li, G.L. Cui, L.J. Kong, X.B. Yang, MIMO radar moving target detection against compound-Gaussian clutter. *Circuits Syst. Signal Process.* **33**(6), 1819–1839 (2014)
8. W.L. Melvin, A STAP overview. *IEEE Aerosp. Electron. Syst. Mag.* **19**(1), 19–35 (2004)
9. C.D. Peckham, A.M. Haimovich, T.F. Ayoub, J.S. Goldstein, Reduced-rank STAP performance analysis. *IEEE Trans. Aerosp. Electron. Syst.* **36**(2), 664–676 (2000)
10. I.S. Reed, J.D. Mallett, L.E. Brennan, Rapid convergence rate in adaptive arrays. *IEEE Trans. Aerosp. Electron. Syst.* **10**(6), 853–863 (2007)
11. P. Stoica, Y. Selen, Cyclic minimizers, majorization techniques, and the expectation-maximization algorithm: a refresher. *IEEE Signal Process. Mag.* **21**(1), 112–114 (2004)
12. Y.L. Tong, T. Wang, J.X. Wu, Improving EFA-STAP performance using persymmetric covariance matrix estimation. *IEEE Trans. Aerosp. Electron. Syst.* **51**(2), 924–936 (2015)
13. H. Wang, L. Cai, On adaptive spatial-temporal processing for airborne surveillance radar systems. *IEEE Trans. Aerosp. Electron. Syst.* **30**(3), 660–670 (1994)
14. C. Xiang, D.Z. Feng, H. Lv, J. He, H.W. Liu, Three-dimensional reduced-dimension transformation for mimo radar space time adaptive processing. *Signal Process.* **91**(8), 2121–2126 (2011)

15. Z.C. Yang, Y.L. Qin, R.C. Lamare, H.Q. Wang, X. Li, Sparsity-based direct data domain space–time adaptive processing with intrinsic clutter motion. *Circuits Syst. Signal Process.* **36**(1), 1–28 (2017)
16. B. Zheng, G.S. Liao, R.B. Wu, Y.H. Zhang, Y.L. Wang, 2-D temporal-spatial adaptive clutter suppression for phased array airborne radars. *Acta Electronica Sinica* **21**(9), 1–7 (1993)
17. Y. Zhou, D.Z. Feng, G.H. Zhu, W.K. Nie, The post-Doppler adaptive processing method based on the spatial domain reconstruction. *Signal Process.* **111**(e), 89–93 (2015)
18. Y. Zhou, B. Jiang, W.K. Nie, W.X. Zhang, The space–time processing method based on the spatial decomposition in airborne MIMO radar. *Acta Electronica Sinica* **45**(10), 2348–2354 (2017)